DOCUMENT RESUME

ED 287 291 FL 016 947

AUTHOR Lange, Dale L.; Lowe, Pardee, Jr.
TITLE Grading Reading Passages According to the ACTFL/ETS/ILR Reading Proficiency Standard: Can It Be Learned?
PUB DATE 87
NOTE 22p.; In: Language Testing Research; Selected Papers from the Colloquium (Monterey, California, February 27-28, 1986); see FL 016 938.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Entrance Examinations; *Difficulty Level; Higher Education; Interrater Reliability; *Language Proficiency; *Language Tests; *Rating Scales; Reading Comprehension; Reading Skills; *Reading Tests; Scoring; Second Language Instruction
IDENTIFIERS *ACTFL ETS Language Proficiency Guidelines; Interagency Language Roundtable; *Ranking

ABSTRACT

A study investigated the use of reading proficiency scales developed by the American Council on the Teaching of Foreign Languages (ACTFL), Educational Testing Service (ETS), and Interagency Language Roundtable (ILR) for meaningful rank-ordering and assigning levels of second language competence to reading passages. In a proficiency test writing workshop in which participants were writing items for a potential college entrance standard and graduation test, the participants ranked and rated reading items in several subsets. Results of the ranking task suggest that consensus for ranking exists. Results of the rating task suggest it is possible to match reading passages to suitable ACTFL/ETS/ILR definitions, despite the limited experience of the raters. These results and participant comments indicate that it is possible for potential users to internalize the ACTFL/ETS/ILR standards and apply them accurately to grading passages. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED287291

GRADING READING PASSAGES ACCORDING TO THE ACTFL/ETS/ILR READING PROFICIENCY STANDARD: CAN IT BE LEARNED?[1]

Dale L. Lange
University of Minnesota

Pardee Lowe, Jr.
Central Intelligence Agency

Introduction

In its American Council on the Teaching of Foreign Languages (ACTFL)/Educational Testing Service (ETS) incarnation (1982, rev. 1986), the Interagency Language Roundtable (ILR) Oral Proficiency Scale and its accompanying levels have been successfully transferred to academe. In contrast, the ILR reading proficiency scale, even in its ACTFL/ETS form, has been less well received. Those attempting to apply it for the first time consistently comment that the reading scale seems harder to grasp than the oral one. Moreover, they understand the rationale for using the reading scale less fully. And more importantly, they question both whether reading proficiency test performances in academe can be rated accurately according to the scale and particularly whether the passages, the comprehension of which forms the basis for ratings, can be properly graded for level.

Purpose of the Study

This study investigates the execution of two tasks pivotal to assigning levels to reading passages: ranking and rating. These tasks were specifically chosen to demonstrate that the ACTFL/ETS and ILR (hereafter referred to as AEI) reading scales provide a meaningful basis for rank-ordering and assigning levels of reading competence. Ranking, the easier of the two tasks, is generally not formally separated from rating as a task. It was chosen, however, to demonstrate to the participants of the item writing workshop that general degrees of difficulty are inherent in passages, a view that contrasts with one regarding each passage as unique and unrankable. If it can be demonstrated that passages possess varying degrees of difficulty, it can be seen that rating by generalized categories, such as the ACTFL/ETS Guidelines, becomes a logical next step. Rating was selected to show that a passage's level of difficulty could usually be matched to a verbal definition without the definition describing every aspect of a passage.

---

[1] We wish to thank Martha Herzog, John Lett, and Ray T. Clifford at the Defense Language Institute for reading an earlier draft of the paper. Their comments have been invaluable. To Martha Herzog and her colleagues at the Defense Language Institute also our gratitude for the excellent selection of texts and for providing insightful commentary on their use in face-to-face reading proficiency interviews.

FL016947

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Schoffman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The Context

The setting for this study was a proficiency test writing workshop at the University of Minnesota in the Summer of 1985 as part of the University's Foreign Language Project. The project focuses on a new language requirement for entrance to the College of Liberal Arts and for graduation with a Bachelor of Arts degree from the college. Arendt, Lange, and Wakefield (1986) have described the project in detail. The new language requirement demands functional competence in a second language in listening, reading, writing, and speaking as opposed to seat time. It is based upon the AEI proficiency statements.[2] (See the attached Guidelines for the generic descriptions.)

A series of steps in the form of three workshops was organized to develop the testing program for the new language requirement. In the first one, participants from the University language departments, the community colleges, private liberal arts colleges, the state university system, and public schools established the expected levels of proficiency in listening, reading, speaking, and writing for both an entrance standard and a graduation requirement. The chosen levels, which follow, can be interpreted with the Guidelines presented in the appendix:

	Entrance Standard	**Exit Requirement**
Listening	Intermediate Low	Intermediate High
Reading	Intermediate Low	Intermediate High
Speaking	Novice High	Intermediate Mid
Writing	Novice High	Intermediate Mid

A detailed statement of expectations for each of the modalities was also constructed. Called a "functional trisection," each statement contains descriptions of the content, functions, and accuracy to be demonstrated by students with each modality for both the entrance standard and graduation requirement.

The second workshop concentrated on the testing of the four language modalities. Participants used the functional trisection they developed as the basis for the discussion of test items and test types, their limitations, and test constraints. They examined multiple-choice and true false items, cloze tests, and such variables as time and facilities for the testing of 1500 students.

---

[2] ACTFL/ETS/ILR (AEI) designates those aspects common to the two scales, such as the AEI concept of "proficiency" (details in Lowe [1986]). The term is not applied to those aspects which differ, such as testing procedures. For example, FSI employs a reading interview while DLI uses a multiple-choice paper-and-pencil testing instrument. In such cases, the procedure is designated by the individual user.

In a third workshop, one week in length, participants wrote a bank of items for the potential entrance standard and graduation tests in listening, reading, speaking, and writing. It was in this latter context that the study took place.

Procedures

Preparation

Although they were already familiar with the AEI scales in all modalities from prior workshops, the first step for this third workshop group was the review of the level definitions for Reading Guidelines on the first day. Further, there was a general discussion of factors contributing to the levels from the grading of a sample of some 27 reading passages by the ILR Testing Committee. The discussion helped participants focus their comprehension of the system on the tasks. After this discussion, the initial ranking and rating of eleven sample ILR texts took place. There was subsequently no discussion of the texts until the two tasks of rating and ranking had been accomplished. Once the ranking and the assignment of ratings were complete, a lively discussion about passage levels and their difficulty arose. This experience served as the basis for the preparation of items for listening, reading, speaking, and writing for the remainder of the week. The discussions helped reinforce the guidelines as a standard for their work.

The Passages

The texts range in difficulty from 0 + /Novice High to Distinguished or Level 5 proficiency descriptions. Due to space limitations, here we will discuss only three representative texts.

Text One, "TV."

The 0 + text, labeled "TV," is a picture of a man carrying a TV set. He appears to be entering a TV repair shop. In the window, there are several signs: TV, Service, Closed Wednesdays. The text contains isolated high frequency words, supported by considerable visual context.

Text Two, "Second Man Held."

2nd Man Held in Twins' Deaths

A second suspect was arrested yesterday in the shooting deaths Friday of Richard F. and Ronald F. Grey, 27-year-old twins, Brockton County police reported.

Samuel K. Cummings, 29, of no fixed address, surrendered at 4 p.m., police said. He was held without bond in the county jail on two first-degree murder charges.

The bodies of the Grey brothers, who lived on Freeville Road SE, were found in separate locations in Upper Boonsboro about 10 miles from the place where they were probably shot -- the rear of the '81' Club in Hampton Heights. Early Sunday, police charged Ralph P. Lucas, 26, of Hampton Heights with two counts of murder.

This is a text of Level 2 difficulty. It is basically factual, but the writer makes a number of assumptions about what readers will understand. For example, the "second suspect" is mentioned before indicating there was a first suspect. And "...bodies...were found in separate locations" is confusing because the victims are twins who presumably lived at the same address. Moreover, the reporter assumes the reader possesses the background knowledge of the legal system and an understanding of language to register the tentativeness of his statements. The text is full of past tense forms. The sentence length increases as does the level of abstraction of language. It appears that compound and complex sentences could play a major role in the comprehension of this text.

Text Three, "When."

> When we look around us today, we see tremendous sums of public and private money poured into artistic and cultural activity at every level. We see a vast network of institutions serving a large and eager but often bewildered public. And, not least, we also see a great deal of unmistakable talent and imagination at work.
>
> Yet how directionless and stymied, how baffled in their purposes, most of this activity and talent seems. In fact, after viewing the art scene all these years, it is impossible for me not to ask: What's wrong here?
>
> Let me put it another way: Why is so much of our art so empty and mean-spirited? Why do so many vaunted reputations turn to ashes so quickly? Why doesn't all the talent, effort, and money produce more of quality and permanence? Why is so much of the criticism lavished upon our art so pusillanimous in confronting failures? And why are our values, tastes, and intellectual loyalties so threadbare? Plainly, there are many things missing in our cultural life. One of the most important of these missing elements, it seems to me, is a critical perspective that is at once serious, high-minded, and disinterested--capable of producing criticism of such integrity that it stands apart from the blizzards of publicity and the unacknowledged social scenarios that today dominate the arts and traduce their objectives.
>
> To put it more bluntly, what is urgently needed in our artistic and cultural life is criticism that asks hard questions, challenges reigning orthodoxies, speaks up for quality and upholds a sense of standards.

This is a Level 4 text. Readers must delve four paragraphs into the text before they discover the topic of the article, "a critical perspective" on the arts. Beginning "in medias res" is a major characteristic of Level 4 texts with their unpredictable chains of thought, the author plunging the reader into the author-made, author-controlled, author-described world. Precise language, synonymous words, phrases, idioms, and constructions mould the reader's perceptions of the author's goal(s). The reader queries also, reading not only between the lines, but reading beyond, forced by the author's unorthodox approach, provocative statements, and careful choice of words to enter a world he does not anticipate. The author achieves his goal when readers emerge somewhere other than where they started.

Tasks and Directions

In the study, eleven new English passages, carefully rank-ordered and graded for level by the ILR Testing Committee, were given blind and in random order to the twenty-five participants for both ranking and rating.

In the first task, ranking, the participants were asked to rank-order the passages for difficulty. There were four rounds to this task. Specifically, the participants received the following instructions:

A. Using your own experience, not those of any possible student test population, rank the passages from easiest to hardest.

B. Assign 1 to the easiest passage, 2 to the second easiest, and so on.

C. Assign 11 to the hardest passage, 10 to the next hardest, and so on.

D. Try to force a choice between any given pair of passages.

E. If two items tie, then write a justification for assigning the same rank to the two passages.

Round One

The purpose of this round was to acquaint the participants with the new passages and to obtain a preliminary overall ranking. Subjects were told to rank all eleven passages and enter their ranking on the Round One reporting sheet (Figure 1 below).

Figure 1

Form Listing of Texts in Random Order

First Name:_____ Last Name:_____

Passage Name	**Level**	**Round**_____
Andres Restaurant	_____	
Citizens Advisory Committee	_____	
Frank O'Hara	_____	
Motor Injured	_____	
OK Mommy	_____	
Second Man Held	_____	
They Call This Progress	_____	
Time To Reconsider	_____	
TV	_____	
When We Look Around	_____	

(Throughout this article, the texts are referred to by the first word in their titles.)

Rounds Two Through Four

The "focus" rounds, Rounds Two through Four, were introduced to lend greater precision to the rankings within any subset of passages, since it was assumed that the participants could discriminate more readily between passages at either end of the continuum than between contiguous ones. Consequently in filling out the reporting sheet for each round, participants were asked to focus in Round Two on the passages they had ranked 1-4, in Round Three on the passages they had numbered 8-11, and in Round Four on the passages they had designated 5-7. They were also told to adjust the rankings of passages other than the ones specifically selected in the round if their focussing mandated a reordering. It was understood that participants might differ in which passages they ranked 1-4, 5-7, and 8-11. In Round Four, participants were instructed to furnish their final overall ranking. We had originally planned to repeat this first task on the last day of the workshop, but the ranking results from the first day suggested that repetition was unwarranted.

In the second task, rating, participants were instructed to assign levels to the passages according to the ACTFL/ETS Guidelines for reading. (See Appendix.) Their choice of levels for each passage was then compared to the levels assigned by the ILR Testing Committee.

To understand how rating was carried out, we need to discuss the relationship of the ACTFL/ETS guidelines to the ILR scale. The relationships of the two systems' reading scales has grown more intricate since the ACTFL/ETS version's last revision. (See Figure 2.)

Figure 2

Comparison of ACTFL/ETS and ILR Scales

ACTFL/ETS SCALE LEVELS	ILR SCALE LEVELS
	5
	4 +
Superior (S, 3-5)	4
	3 +
	3
Advanced Plus (ADV +)	2 +
Advanced (ADV)	2
Intermediate High (IH)	1 +
Intermediate Mid (IM)	
Intermediate Low (IL)	1
NOVICE HIGH (NH)	0 +
NOVICE MID (NM)	
NOVICE LOW (NL)	0
– ABSOLUTE ZERO	

Like the ACTFL/ETS oral scales, the ACTFL/ETS reading scales subdivide the lowest ILR Levels 0 and 1 and assign them verbal descriptions: Novice and Intermediate. Unlike the ACTFL/ETS oral scales, however, the revised ACTFL/ETS reading scales provide numerical base level descriptions (ILR 3, 4, and 5, but not their plus levels) for what in the oral scale corresponds to the omnibus designation, Superior. The interrelationship of the ACTFL/ETS guidelines to the ILR scales became important to the study because some of the ILR sample reading passages required further subdivision (3, 3+, 4, 4+, and 5). (The ACTFL/ETS scale does not include the "plus" level distinctions.) Due to this confusion in the first round, we report only overall results. In the second round, the ILR distinctions were made. The details are given below.

In the rating task, participants were told to match the passage and the linguistic behaviors needed to comprehend it to the Guideline's level descriptions. They were specifically directed to:

A. Match the passage to a single ACTFL/ETS level description.

B. Bracket the passage with a description either side of the one you originally chose. (The purpose of bracketing is to provide three definitions for comparison to ascertain which definition best describes the passage and the behaviors a reader would have to employ to understand the passage.)

C. Determine which of the three descriptions most accurately reflects the nature of the passage and its level of language.

D. If a description to either side of your original choice seems to fit better, bracket again.

E. Repeat the process until you have made your final determination.

Such a procedure generally permits re-rating according to the ILR system except at the higher levels, where the ILR system makes more distinctions than the ACTFL/ETS system, as depicted above.

Results and Discussion

Ranking

Round One revealed that, on average, the participants accurately ranked five of the eleven passages, if one uses the ILR Testing Committee designations as the accuracy criteria. This round identified the two anchor passages, "TV" at the lower end and "Frank" at the higher end. While the rankings of individual participants sometimes varied widely, the averaged group ratings indicated that participants readily identified the polarity of these two extreme passages.

Round Two, with its focus on the easiest passages revealed eleven correct assignments out of eleven. ILR experience suggests that while it is easier than rating, ranking also requires practice, and consequently, slight variations are permissible. In ILR work, passages are generally ranked and rated by committee,

as were both the passages for illustration and those for the blind rating used in this study.

Round Three, with its concentration on the hardest passages, produced nine correct ratings out of eleven. There were misassignments of two passages, "OK" and "Poetry." Round Four, which concentrated on the passages ranked in the middle of the range and provided a final overall rating, also produced nine correct assignments out of eleven. Obviously, with full agreement between the ILR ranking and the participants' average overall ranking in Round Two, the study could have concluded at that point. However, the last two rounds introduced slight variation for four passages: "OK" "Poetry," "They" and "Time."

Results for Part One demonstrate that a consensual basis exists for regarding some passages as harder than others and for ranking the passages accordingly. At this juncture, no overt comparison to the ACTFL/ETS Guidelines took place. Table 1 displays the comparative ranking for each round. Obtaining complete agreement was not the study's major goal, but a high degree of agreement was desirable, and it was achieved.

Table 1

Consensual Ranking of Reading Passages

Passage	Round				UMinn Average	ILR Ranking	ILR Level
	1	2	3	4			
Andres	3	4	4	4	4	4	1/1 +
Citizens	5	5	5	5	5	5	1 +
Frank	11	11	11	11	11	11	5
Motor	6	6	6	6	6	6	2
OK	2	2	2.5	3	2.4	2	1
Poetry	4	3	2.5	2	2.6	3	1
Second	7	7	7	7	7	7	2 diff
They	9	8	8	9	8.3	9	3
Time	8	9	9	8	8.8	8	2 +
TV	1	1	1	1	1	1	0 +
When	10	10	10	10	10	10	4 low

(In round 1 all the texts were ranked. Rounds 2 and 3 dealt with the easiest and hardest texts respectively. In round 4 the midmost texts and then all texts were ranked. Twenty-four participants completed the training.)

Rating

The eleven level assignments in Round One revealed seven complete agreements, three within a plus level of the ILR, and one within a level and a plus of the ILR. An inherent problem exists in assigning levels to passages, as Child (1986) has indicated. The ILR system is designed to rate the processing the reader undertakes, not the product, as in an oral recall protocol in FSI's reading interview. In the FSI procedure, the test taker is given a target language passage to read silently and then is asked to produce a gist. The ILR descriptions, like the ACTFL/ETS Guidelines, are expressed in terms of what a non-native can consistently and sustainedly comprehend. Consequently, to assess the candidate's performance, test administrators must grade each passage so that they can determine whether the candidate understood at the requisite level. This accurate assigning of a level to reading passages obviously requires practice, as does accurate rating of oral interviews.

In Round Two, the participants, on average, assigned the same (or an equivalent ranking) as the ILR Testing Committee did for ten of the eleven passages. In this round, those respondents writing "S" (ACTFL/ETS Superior, which subsumes ILR levels 3, 3 + , 4, 4 + , and 5) were asked to further define that designation, according to the ILR scale. This additional step increased the rate of correct assignments from seven in Round One to ten in Round Two.

Accurate rating depends on internalization of the standard. The fact that the internalization of the scale at the level of the individual is less advanced is shown in this study by the mean score on individual rater/ranker performances, which ranged from a total 3 to 10 out of 11, the mean being 6.5. Passages assigned a split rating, e.g., IM/H, were counted as being appropriately rated if one of the ratings matched that assigned to the passage by the ILR Testing Committee. This method of scoring, requiring the exact original rating or "exact scoring" (ESC) masks the important fact that participants usually rated within a plus point of the ILR Testing Committee's rating.

In rating oral interviews, another method of scoring is applied: in any group of 10 interviews, it is expected that the majority will receive the same ratings as those assigned by experienced testers, with a few deviant scores, no more than two or three within a plus point of an experienced tester's rating. Applying the same approach to the performances of novice passage raters, participants in the present study, scoring 8 to 10, qualify as "proficient" passage raters. Ten participants out of 24 qualified in that range. The ILR Testing Committee, however, has long recognized the difficulty in assigning levels to passages and regularly encourages testers to rate passages in groups. One could also apply a less strict method of scoring, "Proximate Scoring" (PSC), which recognizes that rater trainees are beginning to internalize the standard when they are within a plus point of the original ILR rating. Such scoring suggests that many individuals are indeed achieving internalization of the standard without having become highly proficient at the task.

Using proximate scoring according to the ILR system, we examined the ratings a second time. In this case, the scores ranged from 6-11 with a mean for the group of 9.9. And, although the data base is small, individual patterns could be discerned when proximate and exact scores were compared. Compared to the original ILR ratings, Subject A (Table 2) tended to underrate the difficulty of passages, assigning levels at least a plus point lower on six of the eleven tests.

Subject W demonstrated the opposite tendency, namely overrating three of the eleven passages, underrating one, and rating the remainder exactly as the ILR Testing Committee had done.

Table 2

Assignment of Levels to Reading Passages

Passage	Subject A	B	C	D	E	F	G	H	I	UMinn Mean	ILR Rating
Andres	IM	I	IH	I	A	NH	NH	I	NH	IM	1
Citizens	IM	IH	IH	IH	IL	I	IH	IH	IL	IH	1+
Frank	5	5	5	5	5	5	5	5	5	5	5
Motor	IM	IH	A	A	IH	I	IH	A	IL	A	2
OK	NH	I	IL	NH	I	NH	NH	N	N	IL	1
Poetry	IL	IH	IM	NH	IH	NH	NH	NH	I	IM	1
Second	A	A	A	A	A	IH	I+	A	IH	A	2
They	A	3	3	3	A+	A	A+	A+	3	3	3
Time	A	A+	A+	A+	A+	A	A	A	A	A+	2+
TV	NM	N	NM	N	N	NL	NM	N	NM	NM	0+
When	A+	4	4	4	4	4	4	4+	4	4	4
Language	Sp	F	Sp	Sp	Sp	G	G	G	F	–	–
ESC Mean	3	8	9	8	6	2	4	5	8	10	11
PSC Mean	8	11	11	11	11	8	11	10	8	11	–

TABLE 2 (Continued)

Assignment of Levels to Reading Passages

Passage	Subject J	K	L	M	N	O	P	Q	R	UMinn Mean	ILR Rating
Andres	I	IH	I	–	I	I	NH	I	IH	IM	1
Citizens	IH	IH	A	–	IH	IH	I	IH	IH	IH	1+
Frank	5	5	5	–	5	5	4+/5	5	5	5	5
Motor	A	A	IH	–	IH	IH	A	3	A	A	2
OK	I	I	NH	–	NH	NH	NM	I	NH	IL	1
Poetry	I	I	NH	–	I	I	NH	I	I	IM	1
Second	A+	A	IH	–	A+	A+	A	3	A	A	2
They	3	3	3	–	3	3	A+	3	3	3	3
Time	3	3	3+	–	A+	A+	A+	A+	A+	A+	2+
TV	NH	NH	N	–	NH	NH	NL	NH	N	NM	0+
When	4	4	4+	–	4+	4+	A+	4	4	4	4
Language	G	G	Sp	F	F	F	F	Sp	Sp	–	–
ESC Mean	10	9	3	–	7	7	4	9	9	10	11
PSC Mean	11	11	11	–	11	11	9	9	11	11	–

Table 2 (Continued)

Assignment of Levels to Reading Passages

Passage	Subject S	T	U	V	W	X	Y	UMinn Mean	ILR Rating
Andres	IM	I	I	IM	I	I	IM	IM	1
Citizens	IH	1+	IH	A	IM/H	IH	IH	IH	1+
Frank	5	5	5	5	5	A+	5	5	5
Motor	A	IH	IH	A	A	I	A	A	2
OK	IM	NH/I	I	IM	I	NH	IL	IL	1
Poetry	IM	NM	IH	IM/H	I	IH	IM	IM	1
Second	A	A+	A	A	A+	I	A	A	2
They	4	4+	3	3+	4+	A	3	3	3
Time	4	A	A+	IH	A	A+	A+	A+	2+
TV	NM	N	N	NM	NH	N	NM	NM	0+
When	4+	4	4	4+	5	A+	4	4	4
= = =

Table 3

Mean Ratings by Subject
Compared to the ILR Mean Ratings with Correlations
between Subject and ILR Ratings for the Eleven Passages

SUBJECT	MEAN FOR THE 11 PASSAGES	ILR MEAN	CORRELATION
A	1.918	2.436	.951
B	2.391	2.436	.967
C	2.418	2.436	.984
D	2.300	2.436	.993
E	2.373	2.436	.870
F	1.918	2.436	.991
G	2.073	2.436	.993
H	2.209	2.436	.977
I	1.991	2.436	.926
J	2.527	2.436	.995
K	2.527	2.436	.984
L	2.300	2.436	.965
M	0.000	2.436	.000
N	2.436	2.436	.974
O	2.436	2.436	.974
P	1.945	2.436	.984
Q	2.618	2.436	.956
R	2.391	2.436	.984
S	2.664	2.436	.998
T	2.373	2.436	.954
U	2.391	2.436	.967
V	2.464	2.436	.930
W	2.645	2.436	.977
X	1.755	2.436	.836
Y	2.373	2.436	.993

Figure 3

Plot of Means of Subject Ratings
Showing Deviation from ILR Mean

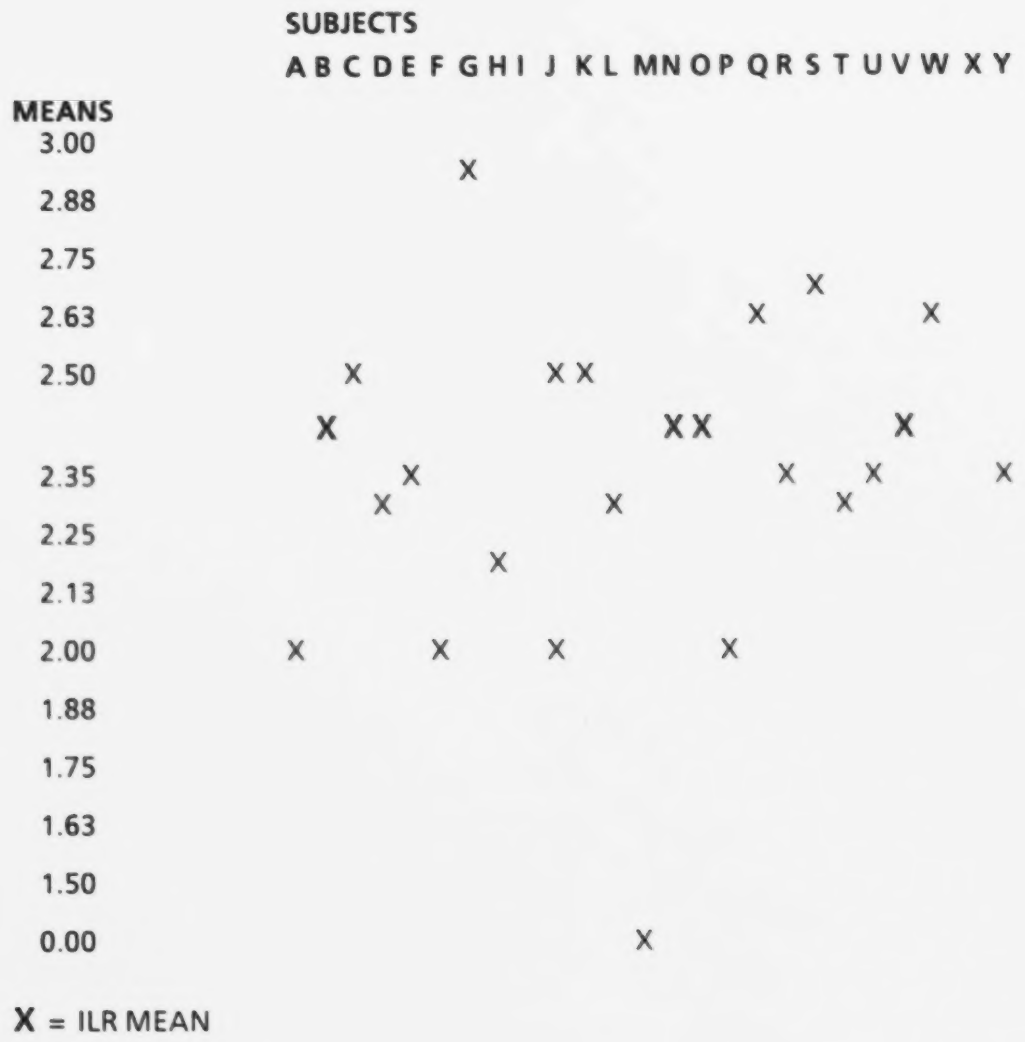

Table 4

Numerical Equivalents
for ACTFL/ETS/ILR Designations

5	=	5.3
4	=	4.3
3	=	3.3
ADV + (2 +)	=	2.8
ADV (2)	=	2.3
IH (1 +)	=	1.8
IM	=	1.3
IL	=	1.1
I	=	1.3
NH	=	0.8
NM	=	0.3
NL	=	0.1
AZ	=	0.0

Figure 3 graphs the direction and extent of deviation to which subjects deviated in their ratings from the ILR. The acceptable range spanned 2.25 to 2.57, raising through Proximate Scoring the number of successful raters to 14 in number. Of the 25 participants, 24 completed the rating tasks. Of those Subject X was clearly off the standard. Five participants graded exactly on standard. Subject G graded most leniently, while Subjects Q, S & W were somewhat lenient. Subjects J and K clustered close to the ILR mean. Subjects A, F, H & P were severe. Subjects D, F, L, R, T, U & Y graded rather conservatively, but within the acceptable range.

Thus Table 4 and the Figure 3 permit a closer assessment of each participant's approximation to the ILR standard. Such variation attests both to the difficulty of the task when undertaken individually as well as to the extent of internalization that proved adequate for the group taken as a whole. Again, the goal of the workshop was not to produce fully trained passage graders, but to impart a sense for the system. The figures presented suggest that this goal was achieved.

Conclusions and Concerns

The outcomes of this study are three-fold. First, the ranking task suggests that it is possible to replicate the overall basis for ranking ILR reading passages from easiest to hardest in an academic setting.

Second, the rating task suggests that it is possible to replicate the AEI rating of such passages, matching the passages to suitable AEI definitions. The workshop did not aim at fully training participants in assigning levels to passages, but rather at imparting to them a sense of a functioning system. Fully training individuals would require greater time on the task.

A third conclusion emerges from the participants' workshop evaluations. Participants consistently remarked that the study with its preparatory phase clearly enabled participants to better understand the nature of the ILR assignment of levels to reading passages for proficiency assessment. Discussing the factors contributing to the rating of the twenty-seven ILR sample passages and then checking the extent of the scale's internalization through the tasks of ranking and rating were particularly helpful.

We believe the results of this study are important for any group that is preparing to write tests based on the AEI proficiency statements for reading. Since our findings in this study parallel our experience with internalizing the oral proficiency standard, namely that one must experience proficiency to use the scales accurately, we hazard the conjecture that such training experiences will prove beneficial in proficiency training for all the skill modalities. As a minimum we recommend that such an exercise should begin every item writing workshop for reading proficiency tests.

Even though it appears possible to rate and rank texts with a fair amount of agreement, we express concerns here when choosing texts either for curricular or evaluation purposes. First, individuals bring their own experiences and meaning to a text when they read. The designation of the level of the text should be considered at best an indication of its proficiency level, particularly with more abstract content at more advanced levels. Every text will be of mixed proficiency levels depending on the amount of stylistic variation in the passage and depending on the world knowledge of the readers (Bernhardt, 1986).

Second, the level given to texts in the process described in this study may not necessarily reflect the competence of the individual responding to them. For example, a level may be assigned to a text, but when responding to testing items on the text, examinees may exhibit understanding of the text above, at, or below the passage's level.

We began with the question: Is it possible for potential users to internalize the AEI standards and apply them accurately to grading passages according to the system? The results of this study strongly suggest that such internalization and application are indeed possible.

References

ACTFL Provisional Proficiency Guidelines. 1982. Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.

ACTFL Proficiency Guidelines. 1986. In Heidi Byrnes and Michael Canale, (eds.), Defining and Developing Proficiency: Guidelines, Implementations, and Concepts. Lincolnwood, IL: National Textbook, pp. 15-24, [Revision of 1982 above.]

Arendt, Jermaine D., Dale L. Lange, and Ray Wakefield. 1986. Strengthening the Language Requirement at the University of Minnesota: An Initial Report. Foreign Language Annals, pp. 149-56.

Bernhardt, Elizabeth B. 1986. Reading in the Foreign Language. In Barbara H. Wing, (ed.), Listening, Reading, Writing: Analysis and Application. Reports of the Northeast Conference on the Teaching of Foreign Languages. Middlebury, VT: The Northeast Conference, pp. 93-115.

Hiple, David V. 1986. A Progress Report on the ACTFL Proficiency Guidelines, 1982-1986. Defining and Developing Proficiency: Guidelines, Implementations, and Concepts. In Heidi Byrnes and Michael Canale, (eds.), Lincolnwood, IL: National Textbook.

Lowe, Pardee Jr. 1986. Proficiency: Panacea, Framework, Process? A Reply to Kramsch, Schulz, and, Particulary, to Bachman and Savignon. Modern Language Journal, 70:391-97.

U.S. Government Interagency Language Roundtable (ILR) Language Skill Level Descriptions: Listening, Speaking, Reading, Writing. Washington, D.C.: Interagency Language Roundtable, 1985. [The descriptions were last revised by Marianne Adams, James Child, John L.D. Clark, Martha Herzog, Pardee Lowe, Jr. with the help of David Hiple, Judith E. Liskin-Gasparro, C. Edward Scebold, and Protase Woodford.]

INTERAGENCY LANGUAGE ROUNDTABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
READING

Preface

The following proficiency level descriptions characterize comprehension of the written language. Each of the six "base level" (coded 00, 10, 20, 30, 40, and 50) implies control of any previous "plus level" designation (coded 06, 16, 26, etc.) will be assigned when proficiency substantially exceeds one base skill level and does not fully meet the criteria for the next "base level." The "plus level" descriptions are therefore supplementary to the "base level" descriptions.

A skill level is assigned to a person through an authorized language examination. Examiners assign a level on a variety of performance criteria exemplified in the descriptive statements. Therefore, the examples given here illustrate, but do not exhaustively describe, either the skills a person may possess or situations in which he/she may function effectively.

Statements describing accuracy refer to typical stages in the development of competence in the most commonly taught languages in formal training programs. In other languages, emerging competence parallels these characterizations, but often with different details.

Unless otherwise specified, the term "native reader" refers to native readers of a standard dialect.

"Well-educated," in the context of these proficiency descriptions, does not necessarily imply formal higher education. However, in cultures where formal higher education is common, the language-use abilities of persons who have had such education is considered the standard. That is, such a person meets contemporary expectations for the formal, careful style of the language, as well as a range of less formal varieties of the language.

In the following descriptions a standard set of text-types is associated with each level. The text-type is generally characterized in each descriptive statement.

The word "read," in the context of these proficiency descriptions, means that the person at a given skill level can thoroughly understand the communicative intent in the text-types described. In the usual case the reader could be expected to make a full representation, through summary, or translation of the text into English.

Other useful operations can be performed on written texts that do not require the ability to "read," as defined above. Examples of such tasks which people of a given skill level may reasonably be expected to perform are provided, when appropriate, in the descriptions.

Reading 0 (No Proficiency)

No practical ability to read the language. Consistently misunderstands or cannot comprehend at all. (Has been coded R-0 in some nonautomated applications.) [Data Code 00]

Reading 0 + (Memorized Proficiency)

Can recognize all the letters in the printed version of an alphabetic system and high-frequency elements of a syllabary or a character system. Able to read some or all of the following: numbers, isolated words and phrases, personal and place names, street signs, office and shop designations; the above often interpreted inaccurately. Unable to read connected prose. (Has been coded R-0 + in some nonautomated applications.) [Data Code 06]

Reading 1 + (Elementary Proficiency, Plus)

Sufficient comprehension to understand simple discourse in printed form for informative social purposes. Can read material such as

announcements of public events, simple prose containing biographical information or narration of events, and straightforward newspaper headlines. Can guess at unfamiliar vocabulary if highly contextualized, but with difficulty in unfamiliar contexts. Can get some main ideas and locate routine information of professional significance in more complex texts. Can follow essential points of written discussion at an elementary level on topics in his/her special professional field.

In commonly taught languages, the individual may not control the structure well. For example, basic grammatical relations are often misinterpreted, and temporal reference may rely primarily on lexical items as time indicators. Has some difficulty with the cohesive factors in discourse, such as matching pronouns with referents. May have to read materials several times for understanding. (Has been coded R-1+ in some nonautomated applications.) [Data Code 16]

Reading 2 (Limited Working Proficiency)

Sufficient comprehension to read simple, authentic written material in a form equivalent to usual printing or typescript on subjects within a familiar context. Able to read with some misunderstandings straightforward, familiar, factual material, but in general insufficiently experienced with the language to draw inferences directly from the linguistic aspects of the text. Can locate and understand the main ideas and details in material written for the general reader. However, persons who have professional knowledge of a subject may be able to summarize or perform sorting and locating tasks with written texts that are well beyond their general proficiency level. The individual can read uncomplicated, but authentic prose on familiar subjects that are normally presented in a predictable sequence which aids the reader in understanding. Texts may include descriptions and narrations in contexts such as news items describing frequently occurring events, simple biographical information, social notices, formulaic business letters, and simple technical material written for the general reader. Generally the prose that can be read by the individual is predominantly in straight forward/high-frequency sentence patters. The individual does not have a broad active vocabulary (that is, which he/she recognizes immediately on sight), but is able to use contextual and real-world cues to understand the text. Characteristically, however, the individual is quite slow in performing such a process. He/she is typically able to answer factual questions about authentic texts of the types described above. (Has been coded R-2 in some nonautomated applications.) [Data Code 20]

Reading 2+ (limited Working Proficiency, Plus)

Sufficient comprehension to understand most factual material in non-technical prose as well as some discussions on concrete topics related to special professional interests. Is markedly more proficient at reading materials on a familiar topic. Is able to separate the main ideas and details from lesser ones and uses that distinction to advance understanding. The individual is able to use linguistic context and real-world knowledge to made sensible guesses about unfamiliar material. Has a broad active reading vocabulary. The individual is able to get the gist of main and subsidiary ideas in texts which could only be read thoroughly by persons with much higher proficiencies. Weaknesses include slowness, uncertainty, inability to discern nuance and/or intentionally disguised meaning. (Has been coded R-2+ in some nonautomated applications.) [Data Code 26]

Reading 3 (General Professional Proficiency)

Able to read within a normal range of speed and with almost complete comprehension a variety of authentic prose material on unfamiliar subjects. Reading ability is not dependent on subject matter knowledge, although it is not expected that the individual can comprehend thoroughly subject matter which is highly dependent on cultural knowledge or which is outside his/her general experience and not accompanied by explanation. Text-types include news stories similar to wire service reports or international news items in major periodicals, routine correspondence, general reports, and technical material in his/her professional field; all of these may include hypothesis, argumentation, and supported opinions. Misreading rare. Almost always able to interpret material correctly, relate ideas, and "read between the lines," (that is, understand the writers' implicit intents in texts of the above types). Can get the gist of more sophisticated texts, but may be unable to detect or understand subtlety and nuance. Rarely has to pause over or reread general vocabulary. However, may experience some difficulty with unusually complex structure and low frequency idioms. (Has been coded R-3 in some nonautomated applications.) [Data Code 30]

Reading 3+ (General Professional Proficiency, Plus)

Can comprehend a variety of styles and forms pertinent to professional needs. Rarely misinterprets such texts or rarely experiences difficulty relating ideas or making inferences. Able to comprehend many sociolinguistic and cultural references. However, may miss some nuances and subtleties. Able to comprehend a considerable range of intentionally complex structures, low frequency idioms, and uncommon connotative intentions; however, accuracy is not complete. The individual is typically able to read with facility, understand, and appreciate contemporary expository, technical, or literary texts which do not rely heavily on slang and unusual idioms. (Has been coded R-3+ in some nonautomated applications.) [Data Code 36]

Reading 4 (Advanced Professional Proficiency)

Able to read fluently and accurately all styles and forms of the language pertinent to professional needs. The individual's experience with the written language is extensive enough that he/she is able to relate inferences in the text to real-world knowledge and understand almost all sociolinguistic and cultural references. Able to "read beyond the lines" (that is, to understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment). Able to read and understand the intent of writers' use of nuance and subtlety. The individual can discern relationships among sophisticated written materials in the context of broad experience. Can follow unpredictable turns of thought readily in, for example, editorial, conjectural, and literary texts in any subject matter area directed to the general reader. Can read essentially all materials in his/her special field, including official and professional documents and correspondence. Recognizes all professionally relevant vocabulary known to the educated non-professional native, although may have some difficulty with slang. Can read reasonably legible handwriting without difficulty. Accuracy is often nearly that of a well-educated native reader. (Has been coded R-4 in some nonautomated applications.) [Data Code 40]

Reading 4+ (Advanced Professional Proficiency, Plus)

Nearly native ability to read and understand extremely difficult or abstract prose, a very wide variety of vocabulary, idioms, colloquialisms, and slang. Strong sensitivity to and understanding of sociolinguistic and

cultural references. Little difficulty in reading less than fully legible handwriting. Broad ability to "read beyond the lines" (that is, to understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment) is nearly that of a well-read or well-educated native reader. Accuracy is close to that of the well-educated native reader, but not equivalent. (Has been coded R-4+ in some nonautomated applications.) [Data Code 46]

Reading 5 (Functionally Native Proficiency)

Reading proficiency is functionally equivalent to that of the well-educated native reader. Can read extremely difficult and abstract prose; for example, general legal and technical as well as highly colloquial writings. Able to read literary texts, typically including contemporary avantgarde prose, poetry, and theatrical writing. Can read classical/archaic forms of literature with the same degree of facility as the well-educated, but non-specialist native. Reads and understands a wide variety of vocabulary and idioms, colloquialisms, slang, and pertinent cultural references. With varying degrees of difficulty, can read all kind of handwritten documents. Accuracy of comprehension is equivalent to that or a well-educated native reader. (Has been coded R-5 in some nonautomated applications.) [Data Code 50]